Torsional path integral Monte Carlo method for calculating the absolute quantum free energy of large molecules

Thomas F. Miller III[a] and David C. Clary[b]
Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

(Received 18 December 2002; accepted 27 February 2003)

A new technique for evaluating the absolute free energy of large molecules is presented. Quantum-mechanical contributions to the intramolecular torsions are included via the torsional path integral Monte Carlo (TPIMC) technique. Importance sampling schemes based on uncoupled free rotors and harmonic oscillators facilitate the use of the TPIMC technique for the direct evaluation of quantum partition functions. Absolute free energies are calculated for the molecules ethane, *n*-butane, *n*-octane, and enkephalin, and quantum contributions are found to be significant. Comparison of the TPIMC technique with the harmonic oscillator approximation and a variational technique is performed for the ethane molecule. For all molecules, the quantum contributions to free energy are found to be significant but slightly smaller than the quantum contributions to internal energy. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1568727]

I. INTRODUCTION

A fundamental objective of modern computational chemistry is the efficient evaluation of free energies. Areas of application for an affordable method of calculating free energy are myriad and span both chemistry and biology. Prominent examples within the realm of biochemistry include the analysis of molecular solvation, protein-ligand binding, and amino-acid sequence dependencies.[1–3]

A multitude of techniques for evaluating both relative and absolute classical free energies have been developed and applied to interesting biochemical systems.[3–10] However, more recent theoretical studies illustrate the significant contributions to chemical equilibrium and dynamics arising from the quantum-mechanical behavior of nuclear degrees of freedom.[11–18] Efforts to incorporate these nuclear quantum effects into free energy calculations have focused primarily on the calculation of reaction rates.[19,20] Less research has been dedicated to the direct evaluation of absolute quantum free energies for equilibrium molecular systems. The largest molecular systems for which absolute quantum free energies have been reported remain limited to only three or four atoms.[21–26]

A recently reported torsional path integral Monte Carlo (TPIMC) technique has made feasible the calculation of quantum statistical expectation values of large molecules at nonzero temperatures.[18] This original report evaluated the internal energy corresponding to the torsional degrees of freedom within various molecules and rigorously included quantum-mechanical contributions. The current paper extends the TPIMC technique to enable the evaluation of absolute quantum free energies. This extension requires the development of a novel importance sampling distribution appropriate for the description of intramolecular torsions.

Section II reviews the torsional path integral Monte Carlo formalism and introduces an importance sampling technique that enables the direct evaluation of molecular partition functions. Section III specifies the details of the reported calculations, and Sec. IV presents a discussion of these results.

II. THEORY

A. Torsional path integral representation

Path integral theory in quantum statistical mechanics derives directly from the representation of the canonical partition function Q as the trace of the equilibrium density operator $\rho = e^{-\beta H}$, where $\beta = 1/kT$.[27,28] The detailed arrival at a computationally useful expression of the canonical partition function of torsional systems has been provided elsewhere, so the result is merely presented here.[18,29,30] For an N torsional system with each torsion exhibiting a moment of inertia I_j, the trace of the density operator ρ may be discretized P times to yield the following approximate relationship for the partition function:

$$Q = \left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} I^{1/2} \prod_{j=1}^{N} \int_0^{2\pi} d\theta_{j,1} \prod_{t=2}^{P} \left[\int_{-\infty}^{\infty} d\theta_{j,t}\right] \times \sum_{n_j=-n_{\max}}^{n_{\max}} e^{-\beta(V^{\text{int}}+V_{\vec{n}}^{\text{ext}}+V_{\vec{n}}^{\text{wind}})}, \tag{1}$$

where

$$V^{\text{int}} = \frac{P}{2\beta^2\hbar^2} \sum_{j=1}^{N} I_j \left[\sum_{t=1}^{P} (\theta_{j,t} - \theta_{j,t+1})^2\right],$$

$$V_{\vec{n}}^{\text{ext}} = \frac{1}{P} \sum_{t=1}^{P} V\left(\vec{\theta}_t - \frac{2\pi\vec{n}(t-1)}{P}\right),$$

$$V_{\vec{n}}^{\text{wind}} = \frac{1}{2\beta^2\hbar^2} \sum_{j=1}^{N} I_j (2\pi n_j)^2, \tag{2}$$

$$I = \prod_{k=1}^{N} I_k, \text{ and } \vec{\theta}_1 = \vec{\theta}_{P+1}.$$

[a]Electronic mail: thomas.miller@chem.ox.ac.uk
[b]Electronic mail: david.clary@chem.ox.ac.uk

The vector $\vec{\theta}_t$ of length N is composed of the $\theta_{j,t}$ angles corresponding to a particular discretization slice t. Using the language of the *classical isomorphism*, the integer P shall be referred to as the number of Trotter beads.[31] Equations (1) and (2) comprise the uncoupled winding number representation of the canonical partition function and become exact in the limits $P\rightarrow\infty$ and $n_{\max}\rightarrow\infty$ for torsions uncoupled by the kinetic-energy operator.[29,30] The classical result is recovered by setting $P=1$ and $n_{\max}=0$. Equation (1) neglects the off-diagonal terms in the kinetic-energy operator, an assumption expected to be legitimate for molecular systems.[32,33]

Contributions to the partition function corresponding to nonzero uncoupled winding numbers n_j are vanishingly small for molecular torsions simulated at temperatures of biochemical interest.[18,29] This is easily illustrated by considering the winding number term $e^{-\beta V_{\vec{n}}^{\text{winding}}}$ in Eq. (1) for a molecular example. Even for the relatively challenging example of a very light torsion such as a hydroxyl (-OH) group rotating with respect to a fixed molecule at 50 K, contributions to the partition function will decay as $\sim e^{-20n_j^2}$. Therefore, for molecular torsions, the maximum uncoupled winding number $n_{\max}$ may be safely set to zero, substantially simplifying the path integral representation of the partition function in Eqs. (1) and (2).

B. HO/FR distributions for the importance sampling of torsions

The canonical ensemble expectation value of the internal energy M can be expressed in the discrete form

$$\langle E\rangle = -\frac{\partial}{\partial\beta}\ln Q = \frac{\int_i E_i e^{-\beta E_i}}{\Sigma e^{-\beta E_i}}, \tag{3}$$

where $Q=\Sigma_i e^{-\beta E_i}$ is the canonical partition function. Because the Boltzmann weights $e^{-\beta E_i}$ appear in both the numerator and denominator of the final expression in Eq. (3), it is convenient to perform importance sampling according to the normalized Boltzmann distribution $\phi = e^{-\beta E_i}/Q$.[34–36] Doing so yields the reduced variance expression for the average

$$\langle E\rangle_n = \frac{1}{n}\sum_{i=1}^{n} E_i\,. \tag{4}$$

The absolute internal free energy is related to the canonical partition function via

$$A = -\frac{1}{\beta}\ln(Q). \tag{5}$$

Free energy follows directly from explicit evaluation of the partition function. Unfortunately, by sampling states according to the normalized Boltzmann distribution, one fails to obtain a useful expression for Q, as illustrated below

$$\langle Q\rangle_n = \frac{1}{n}\sum_{i=1}^{n}\frac{e^{-\beta E_i}}{\phi} = \frac{Q}{n}\sum_{i=1}^{n}\frac{e^{-\beta E_i}}{e^{-\beta E_i}} = Q. \tag{6}$$

However, if importance sampling is performed with the probability distribution ϕ^X of a model system corresponding to an exactly soluble partition function Q^X, such as the free-rotor or harmonic oscillator partition functions, a more useful relationship is obtained

$$\langle Q\rangle_n = \frac{1}{n}\sum_{i=1}^{n}\frac{e^{-\beta E_i}}{\phi^X} = \frac{Q^X}{n}\sum_{i=1}^{n}\frac{e^{-\beta E_i}}{e^{-\beta E_i^X}}$$

$$= \frac{Q^X}{n}\sum_{i=1}^{n} e^{-\beta(E_i - E_i^X)}. \tag{7}$$

This idea forms the basis for the direct calculation of torsional partition functions reported in this study.

We now outline our strategy for directly evaluating the partition function in the uncoupled winding number representation with $n_{\max}$ set to zero. For this case, Eqs. (1) and (2) may be discretized with respect to torsional angle and written in the simplified form

$$\langle Q\rangle = \left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} I^{1/2}\prod_{j=1}^{N}\left[\sum_{\vec{\theta}_j}\right] e^{-\beta(V^{\text{int}}+V^{\text{ext}})}, \tag{8}$$

where summation is over the allowed configuration space of the torsional angles

$$V^{\text{int}} = \frac{P}{2\beta^2\hbar^2}\sum_{j=1}^{N} I_j\left[\sum_{t=1}^{P}(\theta_{j,t}-\theta_{j,t+1})^2\right],$$

and (9)

$$V^{\text{ext}} = \frac{1}{P}\sum_{t=1}^{P} V(\vec{\theta}_t).$$

Suppose that we consider a model system X with an exactly soluble partition function Q_X. The discretized path integral representation of the partition function for the model system would also assume the form of Eq. (8) and (9), except with V^{ext} expressed in terms of the model potential V_X. For the model system

$$V_X^{\text{ext}} = \frac{1}{P}\sum_{t=1}^{P} V_X(\vec{\theta}_t). \tag{10}$$

If we then importance sample according to the normalized path integral Boltzmann distribution of the model system, a computationally useful expression for the partition function of the physical system may be obtained

$$\langle Q\rangle_n = \frac{Q_X}{n}\prod_{j=1}^{N}\left[\sum_{\vec{\theta}_j}\right] e^{-\beta(V^{\text{ext}}-V_X^{\text{ext}})}, \tag{11}$$

for which n is the total number of Monte Carlo samples performed.

We now return to the problem of large molecules constrained to torsional motion. Within the molecular system, some torsions will clearly enjoy more facile motion than others. Torsions corresponding to terminal methyl groups are expected to spin much more freely than those corresponding to the internal backbone of a very large molecule. We, therefore, propose sampling the torsional configuration space of large molecules using a model composed of torsions described by uncoupled free rotors or harmonic oscillators. The

partition function Q_X can then be written as the product of the partition functions for the rotors or oscillators, each of which is soluble. That is

$$Q_X = \prod_{i=1}^{N} Q_\alpha^i, \tag{12}$$

where N is the number of torsions, $\alpha \in \{HO, FR\}$

$$Q_{\mathrm{FR}}^i = 1 + \sum_{m=1}^{\infty} 2e^{-\beta\hbar^2 m^2/2I_i} \quad \text{and} \quad Q_{\mathrm{HO}}^i = \frac{e^{-\beta\hbar\omega_i/2}}{1-e^{-\beta\hbar\omega_i}}. \tag{13}$$

The angular frequency ω_i is determined by an uncoupled harmonic oscillator approximation for the potential well of the ith torsion. The expression for Q_{FR}^i in Eq. (13) is accurately approximated using

$$Q_{\mathrm{FR}}^i = \int_0^\infty 2e^{-\beta\hbar^2 m^2/2I_i} dm = \sqrt{\frac{2\pi I_i}{\beta\hbar^2}}. \tag{14}$$

We shall henceforth refer to the sampling distribution comprised of uncoupled harmonic oscillators (HO) and free rotors (FR) as the HO/FR sampling distribution. A potential strategy for importance sampling in the fashion of Eq. (11) with the utilization of the HO/FR sampling distribution thus emerges.

In our implementation of the new HO/FR sampling distribution, each torsion in the molecular system must be considered with regard to a criterion that determines whether free rotor (FR) sampling or harmonic oscillator (HO) sampling is more appropriate. Although numerous ways of defining this criterion are conceivable, we decided to choose a cutoff for free rotor behavior based on an uncoupled harmonic oscillator analysis of the optimized molecular geometry. At the energy minimum, an uncoupled harmonic oscillator is fit to the local potential of each torsion using finite differences. The variance $\langle \theta^2 \rangle$ of each approximate harmonic oscillator about its minimum is easily calculated according to

$$\langle \theta_i^2 \rangle_{\mathrm{HO}} = \frac{\hbar}{I_i \omega_i}\left(\frac{1}{2} + \frac{e^{-\beta\hbar\omega_i}}{1-e^{-\beta\hbar\omega_i}}\right). \tag{15}$$

Torsions for which the local potential suggests a great deal of variance in the angle are sampled according to a free rotor distribution. A more localized torsional angle is sampled according to the corresponding harmonic oscillator distribution. The appropriate value for this free rotor cutoff is not obvious, so selection of the parameter will be considered in the Sec. IV.

Within the framework of Fourier path integrals, the idea of employing free particle sampling over nonperiodic degrees of freedom has previously been employed for small molecule applications.[21–23,37,38] It should be noted that for the case of nonzero winding numbers, Eq. (11) is easily generalized. Also, because both the free rotor and the harmonic oscillator probability distributions are strictly positive throughout the entire torsional configuration space, the HO/FR sampling technique is, at least in principle, exact within the torsional PIMC framework.

III. CALCULATION DETAILS

Two molecular potentials were utilized in this study. The torsional potential of a simple ethane model was represented by a one-dimensional (1D) function of the torsional angle, $V(\theta) = C[\cos(3\theta)+1]$. The constant $C = 1.33$ kcal mol^{-1} was parametrized to the MM3Pro force field.[39,40] This simple ethane model potential is positive for all torsional angles, exhibits three equivalent minima at $\pm\pi/3$ and π radians, and has a barrier height of 2.66 kcal mol^{-1} between each pair of minima. The larger molecule calculations in Sec. IV B were performed using the all-atom MM3Pro force field as it is implemented in the TINKER molecular mechanics package.[39–41]

In this study, the TPIMC code has been augmented to calculate internal energy using both standard Boltzmann sampling and the new HO/FR sampling technique. Free energies are calculated via explicit evaluation of the canonical partition function using HO/FR sampling. The TPIMC code uses subroutines from the TINKER package to transform between Cartesian and internal coordinates.

We briefly outline the TPIMC parameters generally employed throughout the study. Specific deviations from the recipe are subsequently mentioned. New TPIMC calculations reported in this study were performed in 150 cycles, each initiated from a randomized geometry. For the ethane and butane calculations each individual cycle included 10^5 Monte Carlo steps, but cycle lengths of 10^7 for enkephalin and 10^8 for octane were necessary to obtain satisfactory convergence. For calculations performed using the MM3Pro potential, Monte Carlo samples were taken once every 20 steps to reduce correlation between the samples. Each Monte Carlo step consisted firstly of uniform random shifts of the entire Trotter chains and secondly of uniform random shifts of the individual Trotter beads. Geometry steps were accepted according to the Metropolis algorithm, and bead and chain steps were chosen at the beginning of each calculation to yield an overall acceptance ratio of ~ 0.5.[42] Details of the algorithm by which the chain and bead step size parameters are determined have been previously discussed.[18]

The torsional moment of inertia I_j about a bond between atoms k and l was defined in terms of the moments of inertia, $\tilde{I}_k$ and $\tilde{I}_l$, of two asymmetrical tops that rotate with respect to one another about the bond.[43,44] That is

$$I_j = \frac{\tilde{I}_k \tilde{I}_l}{\tilde{I}_k + \tilde{I}_l}. \tag{16}$$

Each of the two tops correspond to the molecular moieties attached to the atoms terminating the bond.[45] In this study, the torsional moments of inertia are calculated only once and assumed to remain constant throughout the simulation. The moments of inertia are calculated using the starting geometry of the molecule at the global minimum of the potential-energy surface. Although the moments of inertia will, in general, vary as a function of the torsional angle vector, previous work has suggested that such changes have little impact on the statistical averages of very large molecules.[45] Various techniques to update the moments of inertia during the

TPIMC simulation are easily conceived and under investigation. The moments of inertia reported and employed in the previous study have since been corrected to correspond to the molecular applications.[18]

Computational time is dominated by calls to the molecular potential energy subroutine. A benefit of the HO/FR importance sampling scheme is that the computational cost of evaluating the sampling distribution potential is negligible with respect to the cost of evaluating the real molecular potential. Therefore, large numbers of Monte Carlo steps can be performed between samples, and the decorrelation between samples is achieved with no significant increase in computational cost. For the largest TPIMC calculations performed, ~10 days per bead were required on a 1.26 GHz Pentium III processor. The TPIMC method scales linearly with number of beads and number of Monte Carlo steps. For the 4 and 5 bead calculations of octane, the number of calculation cycles performed was reduced from 150 to 75.

Internal energies and free energies obtained with an uncoupled harmonic oscillator approximation are also available for all molecules in this study. For a system of N torsions, the harmonic oscillator approximation to internal energy is determined using

$$\langle E\rangle = \sum_{j=1}^{N} \frac{\hbar\omega_j}{2} + \sum_{j=1}^{N} \frac{\hbar\omega_j e^{-\beta\hbar\omega}}{1-e^{-\beta\hbar\omega}}, \tag{17}$$

where $\omega_j = \sqrt{k_j/I_j}$, and k_j and I_j are, respectively, the force constant and moment of inertia of the jth torsion. Free energies are determined directly from the harmonic oscillator partition function, evaluated according to Eq. (13) to yield

$$\langle A\rangle = \sum_{j=1}^{N} \frac{\hbar\omega_j}{2} + \sum_{j=1}^{N} \frac{1}{\beta}\ln(1-e^{-\beta\hbar\omega}). \tag{18}$$

For the simple ethane model calculated on the sinusoidal potential, variational internal and free energy results were also obtained. By introducing a sufficiently large basis, high numerical accuracy was obtained, so these calculations are assumed to be exact.

IV. RESULTS AND DISCUSSION

A. Simple ethane model

We begin with terrain made familiar by the previous TPIMC study of internal energies.[18] Taking advantage of the simplified, yet realistic, chemical environment afforded by the analytical ethane model, we will analyze the TPIMC internal and free energy results obtained with different sampling distributions. TPIMC calculations are reported using a range of Trotter beads from the classical limit of one to the nearly exact limit of 20. Also reported for the simple ethane models are exact variational and harmonic oscillator approximation results.

Table I shows the ethane internal energies calculated with the TPIMC technique using Boltzmann sampling (BS-TPIMC), free rotor sampling (FR-TPIMC) and harmonic oscillator sampling (HO-TPIMC). Internal energies calculated using the BS-TPIMC technique have been previously published.[18] The FR-TPIMC and HO-TPIMC internal energies replicate the BS-TPIMC results at every Trotter bead. All calculations converged with increasing bead number to the exact variational result of 0.683 kcal mol^{-1}. The internal energy determined with the harmonic approximation to the ethane potential yields an accurate result of 0.648 kcal mol^{-1} at 273.15 K.

TABLE I. Ethane internal energy[a] in kcal mol^{-1} calculated at 273.15 K with the TPIMC technique[b] using various sampling distributions.

P	BS-TPIMC[c]	FR-TPIMC	HO-TPIMC
1	0.5911(3)	0.5912(1)	0.595(2)
2	0.6552(3)	0.6551(2)	0.663(3)
3	0.6708(4)	0.6709(4)	0.680(5)
4	0.6762(6)	0.6768(4)	0.687(6)
5	0.6780(8)	0.6795(8)	0.686(5)
10	0.684(2)	0.680(2)	0.694(3)
15	0.681(4)	0.683(3)	0.689(4)
20	0.691(5)	0.691(5)	0.687(6)

[a]The exact variational and harmonic oscillator approximation results are 0.683 kcal mol^{-1} and 0.648 kcal mol^{-1}, respectively.
[b]Numbers in parentheses indicate standard deviation in the last reported significant digit.
[c]The BS-, FR-, and HO-TPIMC techniques sample the torsional angle according to the molecular, free rotor, and harmonic oscillator Boltzmann distributions, respectively.

Because the torsional angle in ethane rotates relatively easily at standard temperature, the simulation of the sinusoidal ethane potential with a free rotor (or zero) potential is more intuitive than using a harmonic oscillator potential located at the minima of one of the three degenerate ethane potential wells. Therefore, it is not surprising that the internal energy results in Table I illustrate that the FR-PIMC results are more numerically stable than the HO-TPIMC results. Nonetheless, it should be noted that the HO-TPIMC simulation does converge toward the exact result, despite the hindered Monte Carlo sampling distribution. Table II portrays the extent to which the MC simulation overcomes the localization imposed by harmonic oscillator sampling.

In Table II, the first column displays the variance (var $=\langle\theta^2\rangle-\langle\theta\rangle^2=\langle\theta^2\rangle$) of the harmonic distribution actually sampled in the HO-TPIMC simulation. The one-bead result samples the classical harmonic oscillator distribution, and the quantum-mechanical HO distribution is approached with larger bead numbers. This is verified by simple calculation. From classical mechanics we have

TABLE II. Variance in degrees^2 of the harmonic sampling distribution and the simulated torsional angle distribution in the HO-TPIMC calculations of model ethane.

P	HO dist.[a]	Ethane torsion dist.
1	148.7(1)	234(4)
2	168.3(1)	276(6)
3	173.3(1)	299(15)
4	175.3(1)	315(31)
5	175.5(1)	289(9)
10	177.2(1)	290(6)
15	177.5(1)	287(4)
20	177.3(2)	292(9)

[a]The exact quantum and classical harmonic oscillator results are 177.6 deg^2 and 148.8 deg^2, respectively.

TABLE III. Ethane free energy[a] in kcal mol^{-1} calculated at 273.15 K with the TPIMC technique using various sampling distributions.

P	FR-TPIMC	HO-TPIMC	Corrected HO-TPIMC
1	−0.3954(2)	0.2535(5)	−0.3428(5)
2	−0.3668(2)	0.2433(8)	−0.3530(8)
3	−0.3608(2)	0.240(2)	−0.357(2)
4	−0.3583(2)	0.238(3)	−0.359(3)
5	−0.3576(2)	0.240(1)	−0.356(1)
10	−0.3562(2)	0.2391(8)	−0.3572(8)
15	−0.3560(2)	0.2392(6)	−0.3571(6)
20	−0.3559(2)	0.239(1)	−0.357(1)

[a]The exact variational result is −0.3558 kcal mol^{-1}.

$$\langle x^2\rangle_{\text{classical}}=\frac{2\langle V\rangle_{\text{classical}}}{I\omega^2}=\frac{1}{\beta I\omega^2}, \tag{19}$$

where I is the ethane moment of inertia and ω is determined from the harmonic fit to the analytical ethane potential. Inserting the appropriate values into Eq. (19) yields 148.8 deg^2 which is recovered precisely by the one-bead harmonic sampling distribution. Analogously, the quantum HO $\langle x^2\rangle$ expectation value can be analytically determined for comparison. The exact value is found to be 177.6 deg^2 which compares favorably with the HO sampling results obtained with larger bead numbers shown in Table II.

The second column in Table II illustrates the degree to which HO-TPIMC simulation actually samples the ethane potential energy surface. In an ergotic and infinitely long calculation, importance sampling will generate exact MC results for the ethane molecule from the HO sampling distribution by reweighting MC samples according to the relative probability of being produced by the actual molecular distribution versus the HO sampling distribution.[46] However, Table II indicates that only the portion of the ethane potential energy surface local to the minimum of the HO potential is actually explored. Even for calculations utilizing large numbers of beads, the RMS value of the torsional angle is only about $\sqrt{292}$ degrees$\approx$17 degrees. With 60 degrees separating the ethane minima from the neighboring barrier maxima and 120 degrees separating ethane minima, there is little hope that more than one of the three ethane minima are being explored during the MC simulation. The strong agreement between the calculated HO-TPIMC internal energies and those determined with BS-TPIMC and FR-TPIMC suggest that the incomplete sampling is not a problem for the internal energy calculation of ethane. However, the problem of incomplete sampling bodes ill for the use of HO-TPIMC in the calculation of quantities requiring a more thorough exploration of the torsional potential energy surface. Obvious examples include the calculation of entropy for ethane or the calculation of internal energy for systems with asymmetric minima.

Table III displays the free energy of the ethane model calculated with the use of FR-TPIMC and HO-TPIMC. As was previously discussed in Sec. II, free energy is not available via the BS-TPIMC approach. The ethane free energy results reported in the table were obtained simultaneously with the internal energy results in Table I at practically no additional computational cost. Also mentioned in the table is the exact variational free energy of ethane calculated to be −0.3558 kcal mol^{-1}. The FR-TPIMC results behave as expected. With increasing number of beads, the free energy calculated with the FR-TPIMC approach converges toward the exact result. However, the second column of Table III shows that naive application of the HO-TPIMC approach to free energy calculation in ethane fails. As was previously discussed in connection with Table II, only one of the three equivalent ethane potential wells is explored during the MC simulations for which a HO distribution was sampled. Naturally, over-localization of the torsional angle will give rise to an underestimation of entropy and the consequent overestimation of free energy observed in Table III. With increasing bead number, the HO-TPIMC results improve marginally but remain far from the exact result.

The HO-TPIMC sampling pathology illustrated in the second column of Table II suggests a simple correction to the free energy results. If it is assumed that the potential surface of ethane is comprised of three uncoupled harmonic oscillators and that our MC simulations only explore one of these three identical wells, the free energy results may be corrected by introducing a multiplicative factor of three to the partition function. That is, if $G_{\text{uncorrected}}=-\ln(Q)/\beta$ where Q is the calculated partition function, then $G_{\text{corrected}}=-\ln(3Q)/\beta$. Column three of Table III shows the corrected free energies obtained by introducing this *ad hoc* triple degeneracy.

For the case of the model ethane potential, the assumptions underlying the proposed correction to the HO-TPIMC approach certainly seem to be valid. The corrected free energies converge smoothly to the exact result with increasing bead number. However, it should be stressed that the arbitrary assumption of degeneracy numbers is inadvisable. Such an approach is not easily generalized to more complicated systems and will fail even for ethane once the assumption of uncoupled harmonic wells breaks down at higher temperatures. For torsions corresponding to extremely facile degrees of freedom, like that of ethane, MC sampling according to the free rotor distribution will be more appropriate.

Intramolecular pair radial distribution functions provide a straightforward means of analyzing general molecular structures and their quantum deviations from classical behavior. For a general hydrogen–hydrogen radial distribution function $G_{\text{HH}}(r)$, the expression $G_{\text{HH}}(r)\mathrm{d}r$ represents the number of pairs of hydrogen atoms expected to be separated by a distance of between r and $r+\mathrm{d}r$. Figure 1 shows various hydrogen pair radial distribution functions for a single ethane molecule using the FR-TPIMC technique. The figure includes only hydrogen pairs for which both hydrogen atoms are not directly bonded to the same carbon atom. There are nine such hydrogen–hydrogen pairs within the molecule. Figures 1(a) and 1(b) display the radial distribution at 10 and 100 K, respectively. These results are calculated with 1, 5, and 15 Trotter beads. Figure 1(c) displays the hydrogen pair distribution calculated with 15 beads at a variety of temperatures.

At 10 K, Fig. 1(a) shows that the intramolecular structure is dominated by the staggered configuration of ethane. The peaks at 2.5 and 3.1 Å correspond, respectively, to the *gauche* and *anti* hydrogen pairs. The relatively narrow peaks

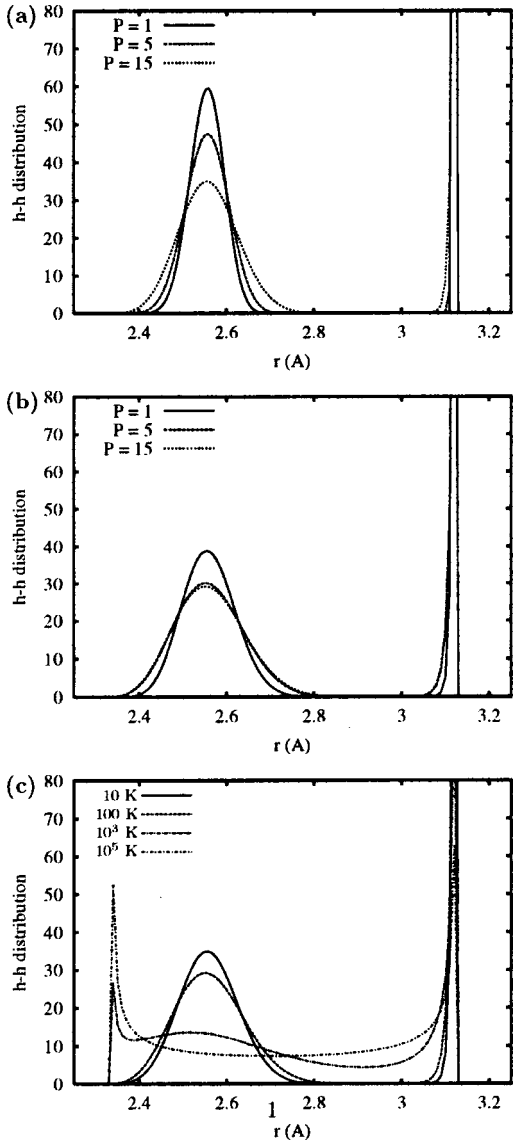

FIG. 1. Hydrogen pair radial distribution functions for ethane at (a) 10 K, (b) 100 K, and (c) various temperatures.

observed in the classical 1 bead distribution broaden for the calculations with 5 and 15 Trotter beads, evidencing the substantial effect of quantum mechanics on the low-temperature ethane structure.

At higher temperatures, the torsional configuration space is more widely accessed by the ethane molecule. In Fig. 1(b), it is seen that the peaks in the distribution function are generally broader than those calculated at 10 K. This expected result for a single ethane molecule differs qualitatively from the radial distribution functions of liquid butane reported by Balog *et al.* for which the decrease of bulk density with temperature gives rise to an anomalous narrowing of the distribution peaks.[14] Figure 1(b) also shows that the impact of quantum mechanics on the intramolecular structure is less important at 100 K than was found at 10 K. The 5 and 15 bead distributions are effectively identical but differ significantly from the classical result.

It is seen in Fig. 1(c) that the preference of the ethane molecule for the staggered configuration erodes with increasing temperature. With higher thermal excitation, the impact of the torsional potential becomes negligible and the distribution peak at 2.5 Å diminishes. At 10^5 K, the radial distribution function is effectively converged to that of a freely spinning classical rotor.

B. Larger molecules

This section reports internal and free energies calculated with the TPIMC technique using HO/FR sampling of the torsional angles. Molecules considered include ethane, *n*-butane, *n*-octane, and the small peptide enkephalin, which have 1, 3, 7, and 33 torsions, respectively. Potential-energy evaluations in this section were performed with the MM3Pro all-atom force field.[39,40]

Table IV reports internal and free energy values calculated for the ethane, *n*-butane, and *n*-octane molecules using the FR-TPIMC technique. From the pit-falls previously observed in calculating free energy for the ethane model using HO sampling, it is concluded that all torsions in the unbranched alkane series are more appropriately sampled using a free rotor distribution. Because the actual MM3PRO atom–atom potential for ethane deviates slightly from the simple ethane model potential described in Sec. III, it is expected that small differences are found between the ethane internal and free energies reported in Table IV and those shown in Tables I and III. The internal energies for all four molecules were previously calculable using the TPIMC technique with Boltzmann sampling (BS-TPIMC).[18] Previously unavailable, however, were the molecular free energies reported in Table IV.

It is seen in Table IV that the behavior of the FP-TPIMC free energy results are qualitatively similar to the internal energies. The quantum contribution to the calculated alkane free energies is somewhat smaller than the corresponding quantum contribution to the internal energies. For ethane, ~ 0.08 kcal mol^{-1} of the internal energy arises from

TABLE IV. Internal and free energies for ethane, *n*-butane, and *n*-octane in kcal mol^{-1} calculated at 273.15 K with the FR-TPIMC technique.

	Internal energy $\langle E \rangle$			Free energy $\langle A \rangle$		
P	Ethane	*n*-butane	*n*-octane	Ethane	*n*-butane	*n*-octane
1	0.5941(5)	1.983(5)	5.04(1)	−0.4139(8)	−1.690(3)	−5.254(3)
2	0.6539(6)	2.080(6)	5.14(1)	−0.3890(7)	−1.647(4)	−5.210(3)
3	0.6689(9)	2.092(7)	5.15(1)	−0.3830(8)	−1.644(3)	−5.201(3)
4	0.671(1)	2.095(9)	5.17(2)	−0.3795(7)	−1.641(4)	−5.200(5)
5	0.677(1)	2.10(1)	5.17(2)	−0.3792(7)	−1.631(3)	−5.199(4)

quantum-mechanical effects, whereas this contribution is reduced by half in the free energy result. Similarly, *n*-butane and *n*-octane exhibit ~ 0.12 kcal mol^{-1} of internal energy and ~ 0.06 kcal mol^{-1} of free energy arising from quantum effects. Note that the uncertainty for the free energy calculations is also much smaller than for the internal energy calculations. As was previously observed in the calculation of internal energies with the BS-TPIMC technique, the majority of the quantum-mechanical effect in the *n*-butane and *n*-octane energies is recovered using only two or three Trotter beads.[18] The ethane molecules, with its considerably lighter moment of inertia, requires a larger number of Trotter beads to converge to a quantum limit.

For the seven-torsion *n*-octane molecule, 1.5×10^{10} MC steps were necessary to obtain satisfactory convergence when free rotor sampling was exclusively employed, as opposed to the 1.5×10^{7} steps sufficient for the three-torsion *n*-butane system. This dramatic increase in the number of required steps with FR sampling distributions suggests that calculation of the 33-torsion enkephalin molecule would be computationally impossible without the selective use of harmonic oscillator (HO) sampling. A systematic approach for deciding which torsions are most appropriately sampled according to a HO distribution was proposed in Sec. II. The set of torsions are ordered according to the variances of related harmonic oscillators, and torsions corresponding to variances above a prescribed cutoff are sampled according to free rotors, whereas torsions below the cutoff are sampled according to the related harmonic oscillator distribution. A suitable value for this cutoff and the number of MC steps needed for convergence are now determined for the enkephalin molecule.

As a reference point, a one-bead enkephalin internal energy calculation was performed using standard BS-TPIMC with 1.5×10^{8} MC steps. It should be noted that this calculation is 75 times longer than the previously published one-bead BS-TPIMC calculation for enkephalin.[18] The value of 17.23(4) kcal mol^{-1} obtained from this calculation serves as the target value for one-bead calculations using HO/FR sampling. Higher-bead BS-TPIMC calculations reported in this study used at least 10^{8} MC steps. Table V shows the dependency of the one-bead enkephalin internal energy calculations on the HO/FR sampling distribution and the number of MC steps.

Table V shows that the HO/FR-TPIMC calculation utilizing nine free rotor distributions and 24 harmonic oscillator sampling distributions (HO/FR=24/9) and 1.5×10^{9} MC steps obtained an adequate value for the one-bead internal energy of enkephalin. It can be seen that employment of fewer FR sampling distributions generally led to underestimation of the internal energy. We assume that a satisfactory simulation of the enkephalin molecule is obtained with the 24/9 sampling distribution and longer MC runs. These parameters are employed in the higher-bead calculations of internal and free energy for enkephalin.

TABLE V. Dependence of one-bead internal energy calculations for enkephalin on HO/FR sampling distributions and number of MC steps.[a]

No. HO/FR	1.5×10^{8} steps	1.5×10^{9} steps
29/4	16.31(8)	16.54(8)
27/6	16.0(1)	15.5(1)
24/9	18.4(1)	17.0(1)

[a]A target value of 17.23(4) kcal mol^{-1} was determined using TPIMC with Boltzmann sampling.

TABLE VI. Enkephalin internal ($\langle E\rangle$) and free ($\langle A\rangle$) energies in kcal/mol calculated at standard temperature.

P	BS-TPIMC $\langle E\rangle$	HO/FR-TPIMC $\langle E\rangle$	HO/FR-TPIMC $\langle A\rangle$
1	17.23(4)	17.0(1)	−38.70(7)
2	17.50(6)	17.5(2)	−38.68(8)
3	17.58(5)	17.4(2)	−38.44(7)
4	17.69(8)	17.2(3)	−38.67(8)

Having determined appropriate parameters for the HO/FR-TPIMC calculation of enkephalin, the internal and free energies are reported in Table VI. The internal energies obtained using HO/FR sampling are in agreement with the BS-TPIMC results. The greatest discrepancy is in the four-bead calculation, for which Monte Carlo convergence is more difficult to fully obtain. In concurrence with previous TPIMC calculations on large molecules, the bulk of the quantum energy is obtained using only two or three Trotter beads. The quantum contribution of ~ 0.25 kcal mol^{-1} to the enkephalin free energy is approximately half of the ~ 0.5 kcal mol^{-1} quantum contribution to internal energy. A similar relationship between quantum free and internal energy was previously noted for the alkane series.

Using the same parameters, it is straightforward to calculate the internal and free energies of enkephalin at 50 K using the HO/FR-TPIMC technique. These results are reported in Table VII and show greater quantum effects than the corresponding calculations at standard temperature. The four bead calculation in Table VII exhibits a quantum contributions of ~ 0.5 kcal mol^{-1} to the free energy and ~ 0.9 kcal mol^{-1} to the internal energy. Although the quantum description of enkephalin at 273.15 K was achieved with only 2 or 3 Trotter beads, the quantum contributions to internal and free energy increased for all four calculations reported in Table VII. Larger quantum effects are found at lower temperatures, but more Trotter beads are needed for full accuracy. This observation is consistent with the temperature dependence of the error term in the Trotter approximation.[47]

C. Quantum contribution to energy

For the molecular calculations previously reported in this study, it was noted that the quantum contribution to internal energy was generally greater than the quantum contri-

TABLE VII. Enkephalin internal ($\langle E\rangle$) and free ($\langle A\rangle$) energies in kcal/mol calculated using the HO/FR-TPIMC technique at 50 K.

P	$\langle E\rangle$	$\langle A\rangle$
1	4.52(7)	−0.16(6)
2	5.09(7)	0.22(5)
3	5.16(8)	0.21(6)
4	5.37(8)	0.36(6)

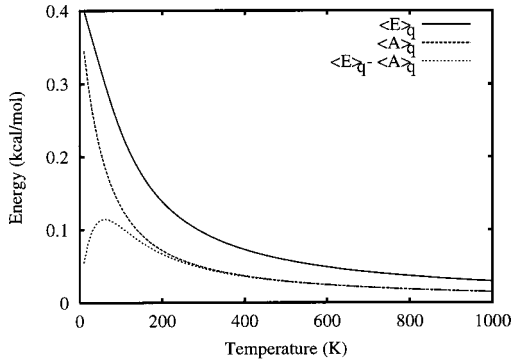

FIG. 2. Quantum contribution to harmonic oscillator internal and free energy.

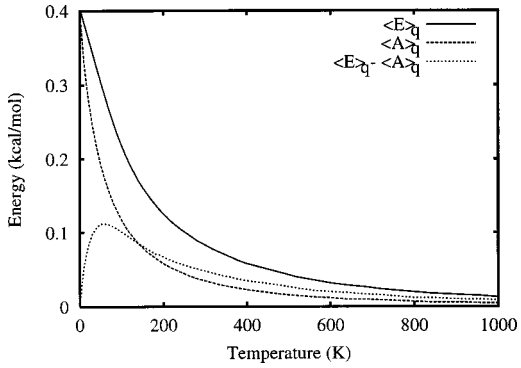

FIG. 3. Quantum contribution to ethane internal and free energy.

bution to free energy by roughly a factor of two. We now consider this trend in greater detail. Define the quantum internal energy $\langle E\rangle_q$ to be the quantum-mechanical contribution to internal energy. That is, $\langle E\rangle_q=\langle E\rangle_{\text{quantum}}-\langle E\rangle_{\text{classical}}$. Likewise define the quantum free energy $\langle A\rangle_q$ and the quantum entropy $\langle S\rangle_q$.

The general observation that quantum free energy is less than quantum internal energy is readily understood. From the standard thermodynamic relationship between free energy, internal energy, and entropy, we easily obtain the equation $\langle A\rangle_q=\langle E\rangle_q-T\langle S\rangle_q$. Because the entropy is related to the volume of configuration space occupied by the molecular system, quantum tunneling is likely to yield a positive $\langle S\rangle_q$ which reduces $\langle A\rangle_q$. However, to explore the possibility of a more specific relationship between quantum internal energy and quantum free energy, the harmonic oscillator and simple ethane model are considered.

For the harmonic oscillator, analytical quantum and classical expressions are readily obtained for both internal and free energy. $\langle E\rangle_{\text{quantum}}$ and $\langle A\rangle_{\text{quantum}}$ were already provided for N uncoupled harmonic oscillators in Eqs. (17) and (18). $\langle E\rangle_{\text{classical}}$ and $\langle A\rangle_{\text{classical}}$ for a single harmonic oscillator are $1/\beta$ and $(1/\beta)\ln(\beta\hbar\omega)$, respectively.

Using these analytical expressions, Fig. 2 presents $\langle E\rangle_q$, $\langle A\rangle_q$, and $\langle E\rangle_q-\langle A\rangle_q$ as a function of temperature for a harmonic oscillator fit to one of the minima of the simple ethane model potential described in Sec. III. In the limit of zero temperature, both $\langle E\rangle_q$ and $\langle A\rangle_q$ converge to the zero point energy of $\hbar\omega/2$. In the high-temperature limit, both converge to zero. However, it is interesting to note that in the high-temperature regime (above 400 K), $\langle E\rangle_q$ is almost exactly twice $\langle A\rangle_q$. This fact is evident from the closeness of the line representing $\langle A\rangle_q$ to the line representing $\langle E\rangle_q-\langle A\rangle_q$. The first-order Taylor expansions for $\langle E\rangle_q$ and $\langle A\rangle_q$ in the high temperature limit $(\beta\rightarrow 0)$ further illustrate this point. For small β, we obtain

$$\langle E\rangle_q=\frac{(\hbar\omega)^2}{12}\beta=\frac{(\hbar\omega)^2}{12kT},$$
$$\langle A\rangle_q=\frac{(\hbar\omega)^2}{24}\beta=\frac{(\hbar\omega)^2}{24kT}=\frac{1}{2}\langle E\rangle_q. \tag{20}$$

The analogous quantities for the simple ethane model potential are presented in Fig. 3. The quantum results in the figure were obtained with the variational technique described in Sec. III, and the classical results were calculated using the FR-TPIMC technique with only a single Trotter bead. Although the general features of Fig. 3 are similar to Fig. 2, it is clear that the quantum free energy does not converge to half of the quantum internal energy for large temperature. The quantum contributions to both free and internal energy drop to zero more rapidly with temperature in the ethane model than is found for the harmonic oscillator. This is due to the asymptotic convergence of the ethane energy levels to those of the free rotor with large thermal excitation. Particularly for anharmonic torsions, a simple relationship between the quantum contribution to internal and free energy is not expected.

V. CONCLUSIONS

In this study, the torsional PIMC technique for the calculation of absolute free energy is introduced and applied to the molecules ethane, *n*-butane, *n*-octane, and enkephalin. The quantum contributions to torsional degrees of freedom are treated in a fashion that converges systematically to the exact result. Utilization of a novel importance sampling distribution based on a model system of uncoupled free rotors and harmonic oscillators enables direct evaluation of the molecular partition function from which free energy is obtained. Although the impact of including the torsional quantum mechanics is generally smaller for free energy than for internal energy, calculations indicate that the quantum contributions to free energy are quite significant. A comparison of the TPIMC technique with the harmonic oscillator approximation and a variational technique is performed for the ethane molecule, and only the TPIMC technique is found to be applicable to large molecular systems. Applications of the TPIMC technique to larger biomolecules and to the analysis of molecular solvation and noncovalent binding are certainly of interest.

ACKNOWLEDGMENTS

One of the authors (T.F.M.) acknowledges fellowships from the Marshall Aid Commemoration Commission and the National Science Foundation. D.C.C. acknowledges a re-

search fellowship from the Leverhulme Trust. The work was supported by the Engineering and Physical Sciences Research Council.

[1] H. Luo and K. Sharp, Natl. Acad. Sci. U.S.A. **99**, 10399 (2002).
[2] M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, Biophys. J. **72**, 1047 (1997).
[3] P. A. Kollman, Chem. Rev. **93**, 2395 (1993).
[4] J. P. M. Postma, H. J. C. Berendsen, and J. R. Haak, Faraday Symp. Chem. Soc. **17**, 55 (1982).
[5] A. Warshel, J. Phys. Chem. **86**, 2218 (1982).
[6] W. L. Jorgensen and C. Ravimohan, J. Chem. Phys. **83**, 3050 (1985).
[7] P. A. Bash, U. C. Singh, R. Langridge, and P. A. Kollman, Science **236**, 564 (1987).
[8] B. Widom, J. Chem. Phys. **39**, 2808 (1963).
[9] R. L. Coldwell, J. P. Henry, and C.-W. Woo, Phys. Rev. A **10**, 897 (1974).
[10] R. Kulver, J. Comput. Chem. **11**, 511 (1990).
[11] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature (London) **397**, 601 (1999).
[12] M. E. Tuckerman, D. Marx, and M. Parrinello, Nature (London) **417**, 925 (2002).
[13] M. Diraison, G. J. Martyna, and M. E. Tuckerman, J. Chem. Phys. **111**, 1096 (1999).
[14] E. Balog, A. L. Hughes, and G. J. Martyna, J. Chem. Phys. **112**, 870 (2000).
[15] C. Alhambra, J. C. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar, J. Phys. Chem. B **105**, 11326 (2001).
[16] S. F. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925 (2001).
[17] C. Alhambra, J. Gao, J. C. Corchado, J. Villa, and D. G. Truhlar, J. Am. Chem. Soc. **121**, 2253 (1999).
[18] T. F. Miller III and D. C. Clary, J. Chem. Phys. **116**, 8262 (2002).
[19] J. Gao and D. G. Truhlar, Annu. Rev. Phys. Chem. **53**, 467 (2002).
[20] L. W. Ungar, M. D. Newton, and G. A. Voth, J. Phys. Chem. B **103**, 7367 (1999).
[21] R. Q. Topper, Adv. Chem. Phys. **105**, 117 (1999).
[22] R. Q. Topper and D. G. Truhlar, J. Chem. Phys. **97**, 3647 (1992).
[23] R. Q. Topper, G. J. Tawa, and D. G. Truhlar, J. Chem. Phys. **97**, 3668 (1992).
[24] S. L. Mielke, J. Srinivasan, and D. G. Truhlar, J. Chem. Phys. **112**, 8758 (2000).
[25] J. Srinivasan, Y. L. Volobuev, S. L. Mielke, and D. G. Truhlar, Comput. Phys. Commun. **128**, 446 (2000).
[26] F. V. Prudente and A. J. C. Varandas, J. Phys. Chem. A **106**, 6193 (2002).
[27] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[28] R. P. Feynman, *Statistical Mechanics* (Benjamin, New York, 1972).
[29] J. Cao, Phys. Rev. E **49**, 882 (1994).
[30] D. Marx and M. H. Muser, J. Phys.: Condens. Matter **11**, R117 (1999).
[31] D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).
[32] A. Vivier-Bunge, V. H. Uc, and Y. G. Smeyers, J. Chem. Phys. **109**, 2279 (1998).
[33] M. A. Harthcock and J. Laane, J. Phys. Chem. **89**, 4231 (1985).
[34] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Methuen, London, 1965).
[35] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Vol. I: Basics* (Wiley, New York, 1986).
[36] M. H. Hansen, W. N. Harwitz, and W. G. Madow, *Sample Survey Methods and Theory* (Wiley, New York, 1953).
[37] R. Q. Topper, Q. Zhang, Y.-P. Liu, and D. G. Truhlar, J. Chem. Phys. **98**, 4991 (1993).
[38] J. K. Hwang, Theor. Chem. Acc. **101**, 359 (1999).
[39] J. H. Lii and N. L. Allinger, J. Am. Chem. Soc. **111**, 8566 (1989).
[40] N. L. Allinger, F. B. Li, and L. Q. Yan, J. Comput. Chem. **11**, 848 (1990).
[41] J. W. Ponder, TINKER *Software Tools for Molecular Design*, *Version 3.8* (2000).
[42] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[43] K. S. Pitzer, J. Chem. Phys. **14**, 239 (1946).
[44] Y. Y. Chuang and D. G. Truhlar, J. Chem. Phys. **112**, 1221 (2000).
[45] D. C. Clary, J. Chem. Phys. **114**, 9725 (2001).
[46] M. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon, New York, 1990).
[47] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).